AN ACOUSTIC STUDY OF THE DEVELOPMENT OF WORD-INITIAL /sP/ CONSONANT CLUSTERS IN THE SPEECH OF A SWEDISH CHILD AGED 1:11 – 2:5 YEARS

Peter E Czigler[1], Jan van Doorn[2] & Kirk P.H. Sullivan[1]
[1]Umeå University, Sweden
[2]The University of Sydney

ABSTRACT: A pilot six-month longitudinal study of the development of the word-initial /s/+plosive cluster was conducted. The experimental participant, a Swedish female child, was 1:11 years of age at the time of the first recording. Homophones were expected for a single word-initial plosive, a word-initial /s/ and a word-initial /s/+plosive cluster. Durational measurements of the plosive and the fricative were made. The observations suggest that the child distinguished between singletons and reduced clusters from 2:1 years of age. The single plosives are produced with aspiration, and the plosives, consistently substituted by the child for consonant clusters, are unaspirated. Further, at the age of 1:11 the fricative sound was very short, then after a unstable period of duration exaggeration, the productions stabilized. This paper has demonstrated that providing the speech material used is well-constructed and can be compared with adult data, the use of duration measurements is an appropriate methodology to follow a child's phonological and phonetic development.

INTRODUCTION

Swedish is considered to belong to the group of languages that are characterised by a relatively complex syllable structure. Both open and closed syllables occur in Swedish. Further, within the same morpheme the phonotactics allow up to three consonants preceding or following the stressed nucleus. In a recent study, Bannert & Czigler (1999) presented a comprehensive quantitative phonetic description of the temporal structure and behaviour of consonant clusters in canonical forms. In addition Bannert & Czigler also described as comprehensively as possible the processes that operate on consonant clusters within words and across word boundaries in spontaneous speech. In their spontaneous speech corpus, which consists of 100 minutes of speech produced by ten speakers, they report 1 643 different types of consonant clusters in 7 419 tokens. Among the most productive consonant clusters, when disregarding word and morpheme position, were the /s/+plosive, (/sP/), and the plosive+/s/, (/Ps/), clusters. Acoustic features of these /sP/ and /Ps/ consonant clusters have been investigated with regard to the timing pattern within the clusters (Czigler, 1998). This study showed that, in accord with studies on other languages (e.g. Klatt, 1976; O'Shaughnessy, 1974; Umeda, 1977), the constituent sounds in Swedish /sP/ and /Ps/ clusters in word final position were of shorter duration than in the singleton context. Additionally, it was found that the duration of the constituent consonant in the first (post-vocalic) position in the cluster was longer than the constituent consonant in the second position. This can be explained by the phonological rule stating that phonologically short vowels are followed by a (phonetically) long consonant even applies in case of postvocalic consonant clusters.

Not only are Swedish children faced with the task of acquiring of a large number of consonant clusters in different position within the word and across word boundaries, they have, in addition, to learn to appropriately apply those phonological rules which make a consonant cluster of four to six (or even more) members pronounceable. Reduction of word-final consonant clusters occurs frequently and systematically in (adult) spoken Swedish. However, word-initial consonant clusters are never reduced (e.g. Gårding, 1978). In an early study, Czigler (1991) examined phonological processes affecting clusters of six and seven consonants in length occurring at word boundaries (CCC(C)#CCC) in the speech of six children aged 4–9 years. In this investigation the children tended to have adult-like reductions in the productions of word-final consonant clusters. However, further research is still required to unequivocally determine how the observed children's reductions should be regarded. Do they demonstrate an acquired ability to apply the appropriate pronunciation rules? Or are they a consequence of the children's developmental difficulties with the timing and motor control of the individual articulatory gestures in complex clusters? The word-final consonant clusters were more likely to be reduced, or were more sensitive to other phonological processes than the word-initial clusters. Simplification processes working in the word-initial consonant clusters occurred more often in the speech of the younger children, i.e. the 4 year olds.

The phonological development of Swedish children is described as a stage model by Nettelbladt (1983). With regard to the development of consonant clusters Nettelbladt places the various developmental phases in her four-stage model as follows. First, homorganic syllable-final two-member consonant clusters can occur at Stage II. At Stage III both syllable-final and syllable initial consonant clusters can occur, although in these traces of harmony restrictions and positional constraints can be detected. At Stage IV two-member syllable-final and syllable-initial consonant clusters are established, and three-member clusters in corresponding positions can occur. According to Nettelbladt there is a sharp jump between Stage II and Stage III, when a number of achievements take place, and the child's phonology is reorganised: "there is a shift in attention from word patterns to syllable shapes and later to segment contrasts" (1983:174).

The differences between children's productions and the adult-like target productions are traditionally described as simplification processes (Ingram, 1974). A number of developmental phonological processes for Swedish children with dysphonology are described and exemplified by Nettelbladt (1983). A brief phonetic description of the development of Swedish consonant clusters was presented by Linell & Jennische (1980). They longitudinally observed two children. In their description they did not include word-medial and word-final consonant clusters; these clusters are of a much greater number and are more complex than word-initial clusters. The word initial clusters are divided into those with an /s/ and those without /s/. Word-initial two-member consonant clusters with /s/ are: /sp, st, sk, sm, sn, sl, sv/, and three member consonant clusters are: /spj, spl, spr, str, skr, skv/. Word-initial two-member consonant clusters without /s/ are: /pl, bl, kl, gl, fl; pr, br, tr, dr, kr, gr, fr, vr; kn, gn, fn; tv, dv, kv, pj, bj, fj, mj, nj/. As a main tendency, Linell and Jennische pointed out that in clusters with /s/ children delete the first consonant of the cluster, while in clusters without /s/ the second consonant is more likely to be deleted. Linell and Jennische paid particular attention to clusters in which /s/ is followed by a voiceless stop /p, t, k/. In Swedish, these voiceless stops in word-initial position are pronounced with a relatively strong aspiration. However, when these stops are preceded by an /s/, they become unaspirated. Linell and Jennische observed the following realisations of word initial /sP/-clusters in the speech of their two subjects:

(1) The /sP/ cluster is substituted by an aspirated voiceless plosive: [p^h, t^h, k^h].

(2) The /sP/ cluster is substituted by an unaspirated voiceless plosive: [p, t, k] – which in turn can be interpreted as /s/-deletion with the unaspirated plosive remaining.

(3) A cluster is produced in which the /s/ is substituted by /h/ followed by an unaspirated plosive: [hp, ht, hk].

(4) In some exceptional cases the plosive is deleted, and the cluster is substituted by an /s/.

It is Interesting to note that, Linell and Jennische did not report the relatively common realisation of the /sP/-cluster as [d]. Linnell and Jennische observed four different ways in which consonant clusters with /s/ + one of following consonants: /m,n,l,v/ were realised:

(5) The /s/-cluster is substituted by a dental plosive or [h]: "snäll" ([snɛlː]) [tɛlː], [dɛlː], [hɛlː].

(6) Deletion of /s/: [nɛlː].

(7) Fusion in the sense that the [-voice] feature of the /s/ influences the following consonant (assimilation): [n̥ɛlː], [^{h}n̥ɛlː]

(8) /s/ is substituted by [h]: [hnɛlː].

(9) In few cases the second consonant is deleted, and the cluster is substituted by an /s/.

As Linell and Jennische point out, (6)-(8) above can even be seen as developmental steps. They also commented that deletion of second consonant in /s/-clusters (see (4) and (9) above) is very uncommon. In relation to the other type of initial two-member consonant clusters, i.e. those without /s/, the following realisations were observed:

(10) Deletion of the first consonant: kniv ([kniːv] knife) [niːv].

(11) Deletion of the second consonant: [kiːv].

(12) Fusion: [n̥iːv].

(13) Vocal epenthesis: [kɪniːv].

The difficulties of distinguishing between, and defining the difference between, for example, deletion of the first member of a consonant cluster (10) and what Linell and Jennische term "fusion" (12) auditorily was discussed by Linell and Jennishe. Interestingly no acoustic studies of consonant clusters in the speech of Swedish children have been reported in the literature as far as the authors are aware. Such an acoustic study would illuminate the issues of definition and identification raised by Linell and Jennische and provide a more subtle description of the stages children pass through in the acquisition of consonant cluster than can be achieved using auditory methods alone. The discovery of any such acoustic subtleties will further provide valuable insight into the nature of children's developing phonological systems. Such a study has been carried out on the acquisition of consonant clusters in Australian English (McLeod, van Doorn & Reed, 1996).

THE AUSTRALIAN STUDY

The aim of the Australian investigation was to look for evidence of a developmental stage in speech production where children acoustically mark the substituted consonant in reduced consonant clusters. The study made acoustic comparisons of homonym pairs produced as a result of cluster reduction by a group of 16 young Australian children aged 2;0 to 2;11 years. Results for temporal measures showed that for word-initial /s/ plus stop clusters which had been reduced to a stop, the mean aspiration duration for the stop in the cluster target word (26.3ms) was significantly less than that for the singleton target word (80.2 ms). The mean duration for /s/ in reduced clusters (201.8ms) was not significantly different from the singleton context (200.4 ms). For the target cluster /kl/ the mean VOT for /k/ was greater (but not statistically significant) for the target cluster context (90.8ms) than for the singleton target context (72.3ms).

THE SWEDISH STUDY

The pilot study presented here is an account of a preliminary acoustic analysis of word-initial potential homophones produced as a result of cluster reduction in a young child's speech. The aim of the investigation was to two-fold. First, to obtain some initial acoustic data on Swedish children's consonant cluster acquisition and compare the acoustically observed changes with those described based on auditive analysis of child speech by Linell and Jennische, and second, to compare the acoustic finding with those of the Australian study (McLeod, van Doorn & Reed 1996).

METHOD

Experimental participant

One child growing up in a monolingual Swedish environment participated in this pilot study. According to an interview with the child's mother, a researcher at the Department of Philosophy and Linguistics, Umeå University, Sweden, the child was judged to be within normal limits of hearing, speech, language and cognition for her age.

Speech corpus

The speech corpus consists of four separate recordings (from 1:11 to 2:5 years of age) of 34 target words. Four sets of three target words were used in this investigation: (1) *pil* (arrow), *sil* (herring), *spis* (stove); (2) *potta* (potty), *socka* (sock), *spocka*; (3) *tumme* (thumb), *sol* (sun), *stol* (chair); (4) *ko* (cow), *sol* (sun) and *sko* (shoe). Each word set contained a word with an initial stop, a word with an initial /s/ and a word with and initial /sP/-consonant cluster. As far as possible words within the child's active vocabulary were chosen. Where there was no suitable word within the child's active vocabulary a word, which was considered to be a suitable approximation to the idea member of the set, was chosen. In one case a nonsense word was created, *spocka*, as the name for a 'new' animal.

Recording procedure

Longitudinal data was recorded at 1:11 and 2:1 years of age, and at 2:4 and 2:5 years of age. At each recording session three sets of 34 target pictures were presented for the child. Words were elicited spontaneously as far as possible. The high-quality digital recordings were made in a sound attenuated recording studio; these were stored as individual computer files with a recording frequency of 16kHz.

Analysis and Measurements

The standard waveform and spectrogram facilities available in ESPS/*waves*™ were used to make the duration measurements. The duration of the plosives was measured from the onset of the release of the plosive to onset of the first appearance of periodicity in the waveform. The duration of the fricative

/s/ was measured from the onset of the simultaneous appearance of high frequency energy in the waveform and wide-band spectrogram to the onset of voicing of the following vowel.

RESULTS

During the period of the study the child did not develop adult-like word-initial consonant clusters; all word-initial consonant clusters were substituted by a (voiceless) plosive. With a single exception, the child clearly made a distinction between word-initial single voiceless plosives and fricatives. The fricative is an /h/ in the first two recording sessions and a /θ/ in the last two recordings. Further, the child did not make distinguish between dental and velar voiceless stops; /k/ was substituted by /t/.

Comparison of singleton plosive and /sP/-cluster realizations

The first set of target words includes a word with word-initial /p/ and a word with word-initial /sp/. Figure 1a shows the mean durations of the prevocalic plosive for the /p/ in the word *pil*. In the word *spis* the /sp/ is substituted with /p/. Its mean durations are presented in Figure 1b.

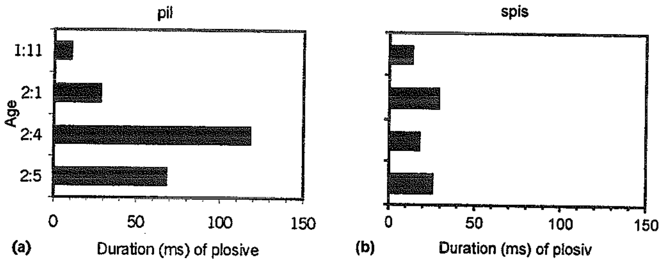

Figure 1. Mean duration (ms) of the plosives (a) in the word *pil*, and (b) *spis* at 1:11, 2:1, 2:4 and 2:5 years of age.

The second set of target words includes a word with word-initial /p/, and a word with word-initial /sp/. Figure 2a shows the mean durations of the prevocalic plosive for the /p/ in the word *potta*. In the word *spocka* the /sp/ is substituted with /p/: mean durations are presented in Figure 2b.

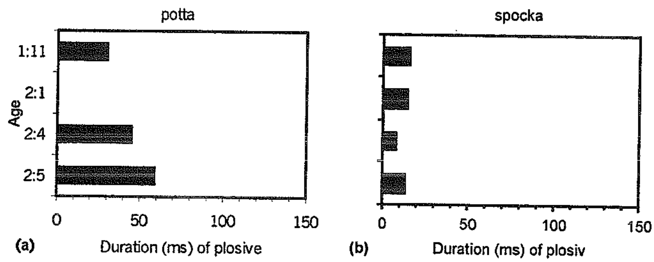

Figure 2. Mean duration (ms) of the plosives (a) in the word *potta*, and (b) *spocka* at 1:11, 2:0, 2:4 and 2:5 years of age.

The third set of target words includes a word with word-initial /t/, and a word with word-initial /st/. Figure 3a shows the mean durations of the prevocalic plosive for the /t/ in the word *tumme*. In the word *stol* the /st/ is substituted with /t/. Its mean durations are presented in Figure 3b. The fourth set of target words includes a word with word-initial /k/, and a word with word-initial /sk/. Figure 4a shows the mean durations of the prevocalic plosive for the /t/ (as /k/-substitute) in the word *ko*. In the word *sko* the /st/ is substituted with /k/: mean durations are presented in Figure 4b.

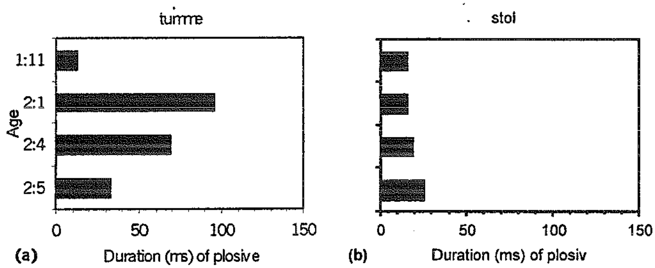

Figure 3. Mean duration (ms) of the plosives (a) in the word *tumme*, and (b) *stol* at 1:11, 2:1, 2:4 and 2:5 years of age.

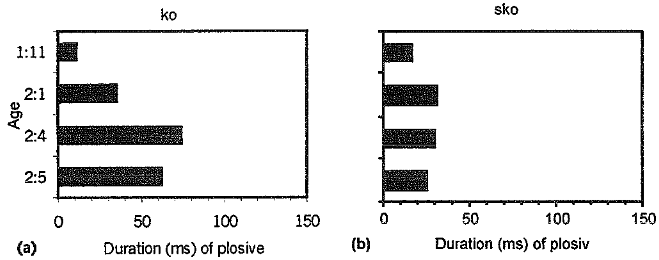

Figure 4. Mean duration (ms) of the plosives (a) in the word *ko*, and (b) *sko* at 1:11, 2:1, 2:4 and 2:5 years of age.

Development of singleton fricative realizations

In each set of words one word with word-initial /s/ occurs. Figure 5a presents the fricative durations in the word *sil*, Figure 5b presents the fricative in *socka*, and Figure 5c in *sol*.

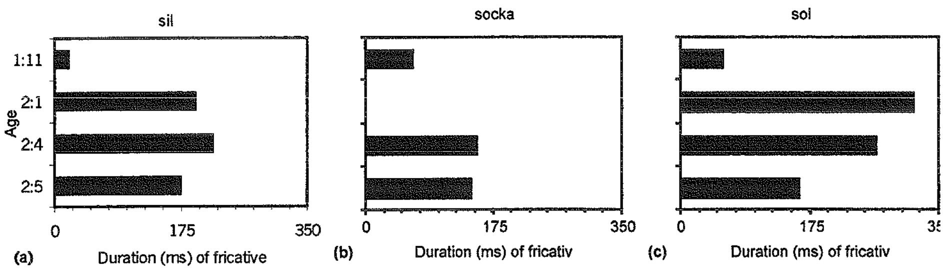

Figure 5. Mean duration (ms) of the fricatives in the words (a) *sil*, (b) *socka*, and (c) *sol*, at 1:11, 2:1, 2:4 and 2:5 years of age.

DISCUSSION

The data presented in this paper is based on one child. Her tokens were not always useable for analysis and thus, there are totally missing data points (see *socka* 2:1, Figure 5), and in other case the presented data may result from anything from a single token up to four tokens. The interpretation of the data is thus qualitative and hints at general trends which demand further more detailed and quantitative investigation with a larger participant base.

Observation of the durational data shows the child distinguished between singletons and reduced clusters from 2:1 years of age. This can be seen in the way the child produced the singletons and the reduced clusters. The single plosives are produced with aspiration, and the plosives, consistently substituted by the child for consonant clusters (/sP/), are unaspirated. From 2:1 years it is clear that the child has acquire the adult phonology which demands that in Swedish voiceless stops are aspirated in word initial position and that these stops are unaspirated after /s/. This finding is consistent with McLeod, van Doorn & Reed (1996). Further, the child gradually decreased the length of single voiceless stop over the period of the study. In order to assess how adult-like the child's productions are at 2:5 years demands analyzed adult data with which the data presented here can be compared. Such data can be found in the VaKoS database (Bannert & Czigler, 1999).

The tendency found for the development of the duration of the aspirated stops can also be seen in the child's single fricative development. At the age of 1:11 the sound was very short, then after a unstable period of exaggeration in duration, the productions stabilized at around 175 ms at 2:5 in all three words. The data also hints that the child by 2:4 is making a distinction between fricative in mono- and bi-syllabic words. Further systematic examination of this topic could throw light on whether children are able to pre-plan their utterances with respect to syllable internal timing in an adult-like manner.

In summary, this paper has demonstrated that providing the speech material used is well-constructed and can be compared with adult data, the use of duration measurements is an appropriate methodology to follow a child's phonological and phonetic development. The inclusion of other acoustic cues such as those suggested by McLeod, van Doorn and Reed (1996) would strengthen this approach.

REFERENCES

Bannert, R. & Czigler, P.E. (1999) Variations in consonant clusters in Standard Swedish, Reports from the Department of Phonetics, Umeå University, PHONUM 7. Text and corpus also available from http://www.ling.umu.se/fon/pub/phonum07.1999.12.html

Czigler, P.E. (1991) Cluster production at grammatical boundaries by Swedish children: some preliminary observations, PERILUS XIII, 73–76.

Czigler, P.E. (1998) Timing in Swedish VC(C)-sequences, Reports from the Department of Phonetics, Umeå University, PHONUM 5.

Gårding, E. (1978) *Talet*, In Pettersson, Å. & Badersten L. Språk i utveckling, Lund: LiberLäromedel, 9–54.

Ingram, D. (1974) *Phonological rules in young children*, Journal of Child Language, 1, 49–64.

Klatt, D.H.(1976) *Linguistics uses of segmental duration in English: Acoustic and perceptual evidence*, Journal of Acoustic Society of America, 59, 1208–1221.

Linell P. & Jennische M. (1980) Barns uttalsutveckling. Lund: LiberLäromedel.

Nettelbladt, U. (1983) Developmental studies of dysphonology in children. Travaux de l'Institute de Linguistique de Lund XIX. Lund: CWK Gleerup.

McLeod, S, van Doorn, J. & Reed, V. (1996) *Homonyms and cluster reduction in the normal development of children's speech*, Sixth International Conference on Speech Science and Technology, 331–336.

O'Shaughnessy, D. (1974) *Consonant durations in clusters*, IEEE Transactions on Acoustics, Speech and Signal Processing, ASSP-22, 282-295.

Umeda, N. (1977) *Consonant duration in American English*, Journal of Acoustic Society of America, 61, 846–858.